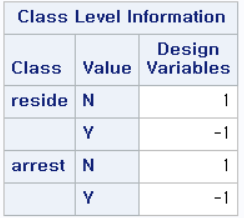
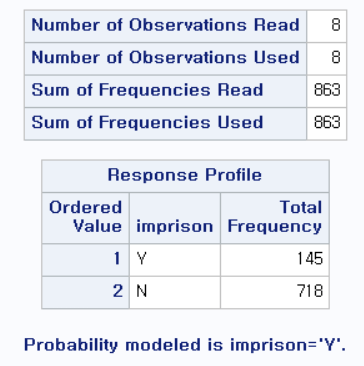
생명과학 데이터분석 과제#2

2014150137 통계학과

박 정진

4-1.

(1)



구속이 ‘예’ 인 경우에 따라 모델링 하였으며,

해당 도시의 거주민인 경우(reside) -1, 아닌 경우 1

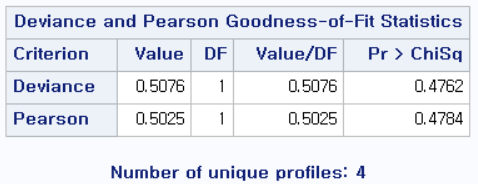
체포 경력이 있을 경우(arrest) -1, 없을 경우 1

* 로짓 모델

( = reside, = arrest)

* 모델 적합성과 회귀계수의 검정

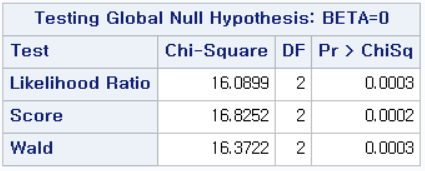
위의 모델은 적합하다. 위의 모델은 적합하지 않다.



Deviance 값이 0.5076으로, P-Value가 0.4762 이다.

즉, 귀무가설을 유의수준 =0.05에서 기각하지 못하므로, 모델은 적합하다고 할 수 있다.

0 Not

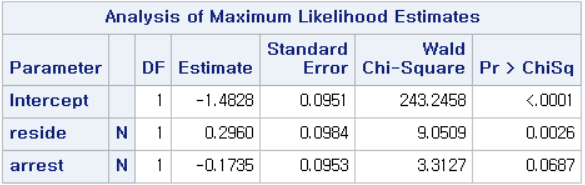


3가지 검정 통계량 모두 p-value가 0.05보다 낮다.

즉, 귀무가설을 유의수준 =0.05에서 기각하므로, 유의한 회귀계수가 존재한다..

* 자료의 모수 추정 결과

각각의 모수,에 대한 추정은 아래와 같다.



각각의 회귀 계수들에 대하여,

의 추정량은 P-value가 0.0026으로, 유의수준 =0.05에서 유의 하다고 할 수 있다.

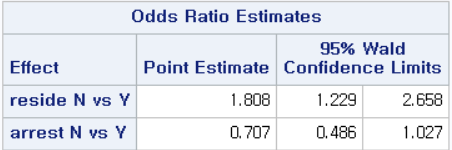
반면, 의 추정량은 P-value가 0.0687으로, 유의수준 =0.05에서 유의하지 않다.

추정치를 바탕으로 모형을 세우면,

( = reside, = arrest)

이 된다.

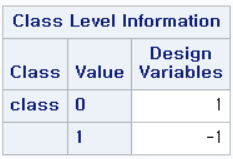
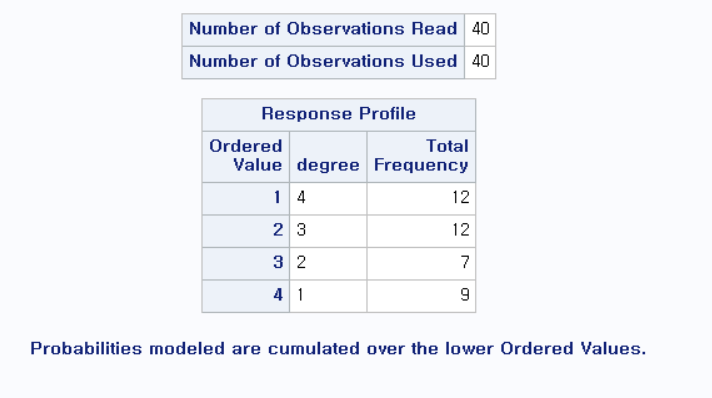
(b)



체포 경력을 고정시켜 놓았을 때, 거주민이 아닌 경우의 오즈는 거주민일 경우의 오즈의=1.808배라고 할 수 있다. 즉, 거주민이 아닌 경우에는 거주민인 경우보다 구속 판결을 받을 오즈가 높다.

거주민 여부를 고정시켜 놓았을 때, 체포 경력이 없는 경우의 오즈는 체포 경력이 있는 경우의 오즈의 =0.707배라고 할 수 있다. 즉, 체포 경력이 없는 경우의 오즈는 체포 경력이 있는 경우보다 구속 판결을 받을 오즈가 높다.

4-5.

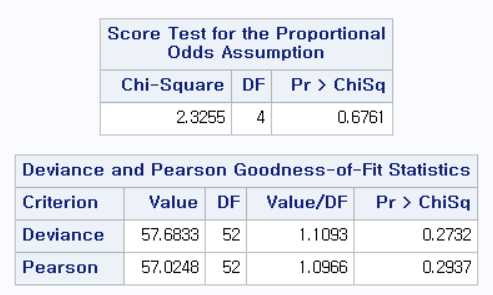


정신 장애의 정도가 낮은 순서에서부터 차례대로 확률이 누적되고 있음을 알 수 있으며,

사회경제적 지위가 0인 경우 1, 사회경제적 지위가 1인 경우 -1이 지정되었다.

사용하려는 누적 로짓 모델은

(j=1, 2, 3) ( 이다.



비례오즈 모델의 경우, 범주에 대해 설명변수들의 효과가 동일함을 가정하기 때문에, 먼저 이것을 검정한다.

비례오즈 가정을 만족한다. Not

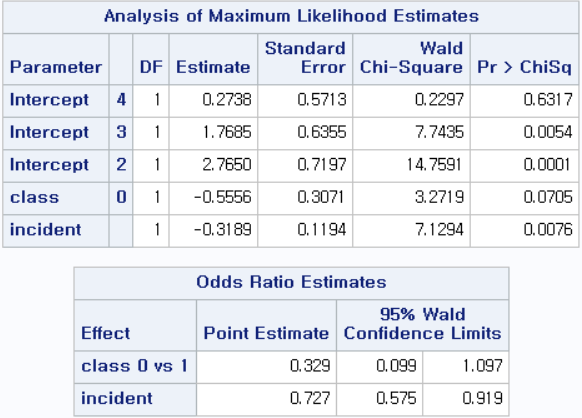
카이스퀘어 통계량의 p-value가 0.6761로, 유의수준 =0.05에서 귀무가설을 기각하지 못한다. 따라서 이 모델은 비례오즈 가정을 만족하고 있다.

가정을 검정한 후에는 모델의 적합성을 평가한다.

모델이 적합하다. Not

Deviance와 Pearson 통계량의 P-value가 모두 유의수준 =0.05에서 귀무가설을 기각하지 못한다. 따라서 위의 모델은 적합하다고 이야기 할 수 있다.

각각의 모수들에 대한 추정은 다음과 같다.



모수 추정치를 바탕으로 각 반응 변수의 수준에 따른 모형식은 다음과 같이 나타난다.

(j=1)

(j=2)

(j=3)

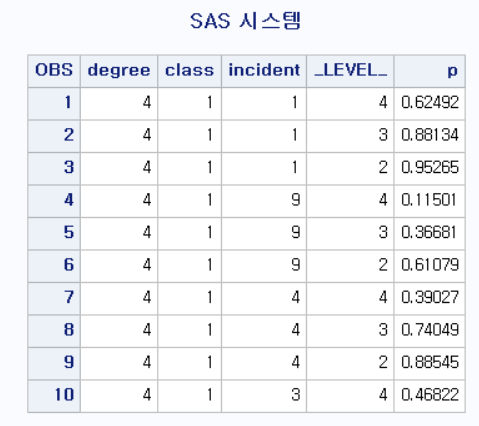
j=1일때부터 정신장애 정도가 점점 높아진다.

사회 경제 지위의 경우, 카이스퀘어 통계량의 P-value가 0.0705로, 유의수준 =0.05에서 유의하지는 않다.

중요한 사건을 고정시켜 놓았을 때, 사회 계층이 0 일때의 오즈는 사회 계층이 1일 때의 오즈의=0.329배라고 할 수 있다.

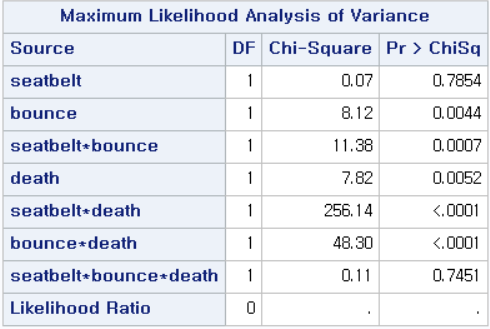
사회 계층이 고정 되었을 때, 중요한 사건을 한 단위 증가시킬 때의 오즈는 이 전의 오즈의 =0.727배라고 할 수 있다.

각 개체의 경우 별 누적확률은 다음과 같다.



즉, class=1, incident=1에서 첫 개체가 약간일 확률은 0.625, 경증일 확률은 0.88134-0.62492, 중증일 확률은 0.95265-0.88134가 된다.

4-7.



X = 안전벨트 착용 여부

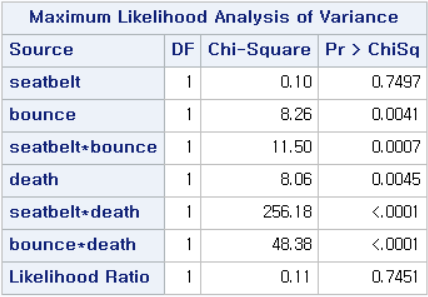
Y = 차에서 튕겨 나감

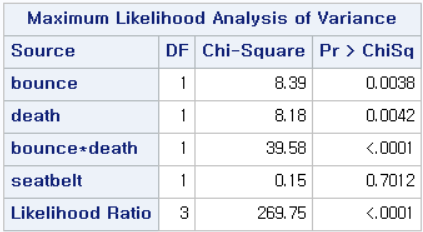
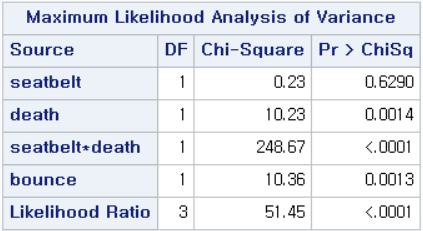
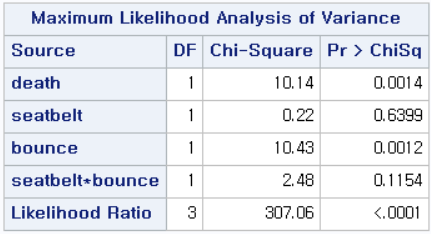
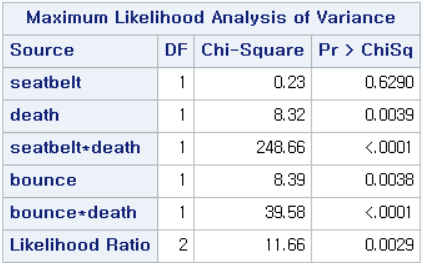
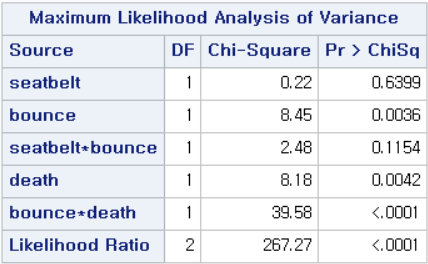
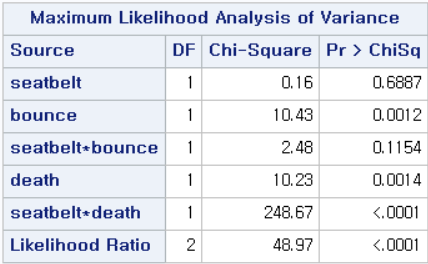
Z = 사망

이라고 할 때,

(XYZ)의 3요인 교호작용을 고려한 모델은 카이스퀘어 통계량이 0.11이고 P-value가 0.7541로, 유의수준 =0.05에서 유의하지 않은 것으로 볼 수 있다. 따라서 2요인 교호작용 까지를 고려한 모델을 살펴보기로 한다.

모형의 간결성 측면에서, (XY, XZ, YZ)를 고려한 모형 외에도 2요인 교호작용 2개의 교호작용 만을 고려한 모형, 1개의 2요인 교호작용을 고려한 모형, 주효과만을 고려한 모형을 모두 적합하여 보았으며, 그 결과값은 아래 OUPUT과 같고 결과를 표로 정리하면 아래와 같다.





|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 로그-선형모형 | 독립유형 |  |  |  |
| (XY,YZ,XZ) | 부분연관모형 | 0.11(1) |  |  |
| (XZ,XY)  (XY,YZ)  (XZ,YZ) | 조건부독립 | 48.97(2)  267.27(2)  11.66(2) | 48.86(1)  267.16(1)  11.55(1) | 3.841 |
| (XY,Z)  (YZ,X)  (XZ,Y) | 결합독립 | 307.06(3)  269.75(3)  51.45(3) | 307.95(2)  269.64(2)  51.34(2) | 5.991 |
| (X,Y,Z) | 상호독립 | 309.54(4) | 309.43(3) | 7.815 |

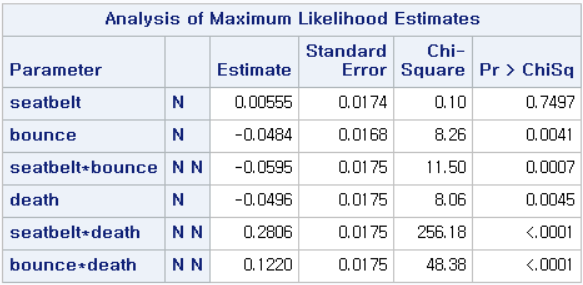
.

.= 우도비 검정 통계량.

=- : 두 모형의 비교를 위한 통계량

4열의 통계량과 5열의 카이스퀘어 통계량을 보았을 때, 다른 모든 모델에 비하여 (XY ,XZ, YZ) 모델이 더 적합함을 알 수 있으며 따라서 이 모형을 선택한다.

즉, 안전벨트-차 튕김-사망 간에는 3요인 교호작용은 없지만 각각의 2개 변수 간에는 교호작용이 모두 존재하고 있다고 할 수 있다.



이에 따르는 모델은, 이 된다.

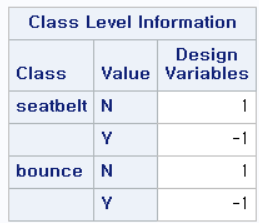
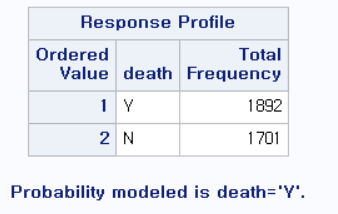
주효과 안전밸트 착용여부를 제외한 나머지 주효과와 2요인 교호작용들은 유의수준 =0.05에서 유의하다고 볼 수 있다.

2요인 교호작용들을 해석해보자면, 안전벨트를 하지 않을 경우 차에서 튕겨져 나가지 않을 확률이 감소한다. 즉, 안전벨트를 하지 않으면 차에서 튕겨져 나갈 확률이 증가한다.

벨트를 착용하지 않은 경우, 죽지 않는 경향이 있다.

차에서 튕겨져 나가지 않는 경우, 죽지 않는 경향이 있다.

(2)

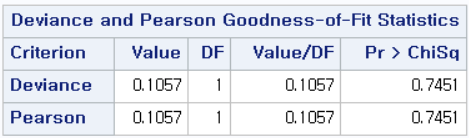


사망여부를 반응변수로 둔 로짓 모형이다.

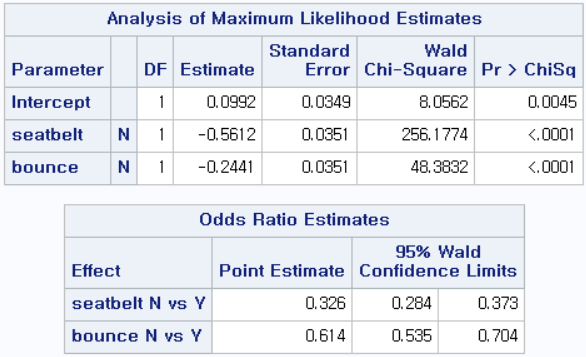
안전벨트 미착용시 1, 착용시 -1

차에서 튕기지 않을시 1, 튕겨져 나갔을 시 -1

의 값이 적용되었다.



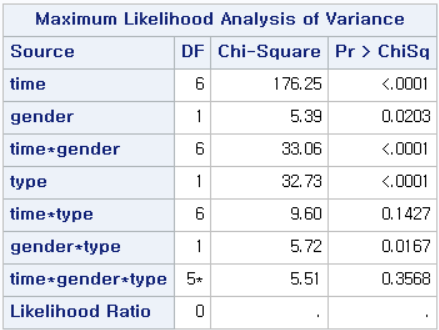
적합도 통계량의 P-값이 모두 0.7451으로 유의수준 =0.05에서 모형이 데이터를 잘 적합한다고 말할 수 있다. 또, 모수추정 결과로 적합된 모형은 아래와 같다.



추정된 오즈비를 통해 안전벨트와 튕김 여부가 죽음에 미치는 영향을 해석할 수 있다. 우선, 차에서의 튕김을 고정시킨 상태에서 벨트를 착용하지 않았을 때의 사망할 오즈는 벨트를 착용 할 때의 사망할 오즈의 배이다. 또, 벨트를 고정시킨 상태에서 차에서 튕겨나지 않았을 때의 사망할 오즈는 차에서 튕겨 나갔을 때의 사망할 오즈의 배 이다.

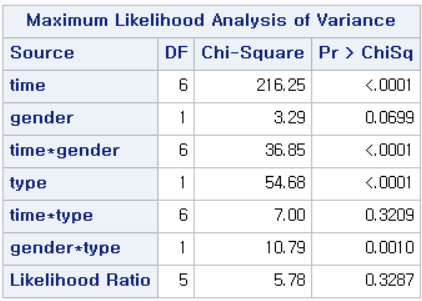
4-9.

(1)

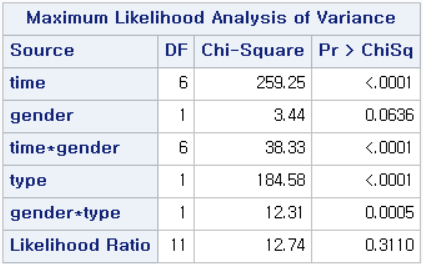
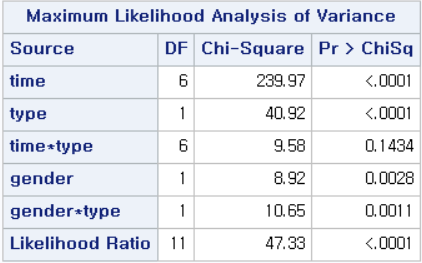
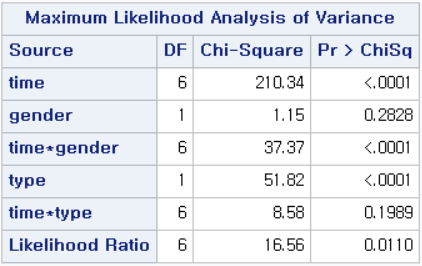


먼저 포화 모형을 고려한다.

3요인 교호작용에 대한 p-value는 0.3568으로 0.05보다 크기 때문에 유의수준 =0.05에서, 3요인 교호작용이 그리 유의하지 못함을 알 수 있다. 따라서 3요인 교호작용을 제외한 부분 독립 모형을 고려한다.



Likelihood Ratio의 카이스퀘어 통계량의 P-value가 0.3287로, 유의수준 =0.05에서 모델이 옳게 적합 되었다고 이야기 할 수 있다. 이것이 최종 모형인 것인가를 알아보기 위해 더 축소된 모델들과 비교를 계속한다.



.

.= 우도비 검정 통계량.

=- : 두 모형의 비교를 위한 통계량

X=시간, Y=성별, Z=고등학교 종류 라고 할 때,

(XY, XZ) 모형의 경우, =16.56-5.78=10.76 > =3.841

이므로, 유의수준 =0.05에서 (XY, XZ) 모델이 더 적합하다고 말하기 힘들다

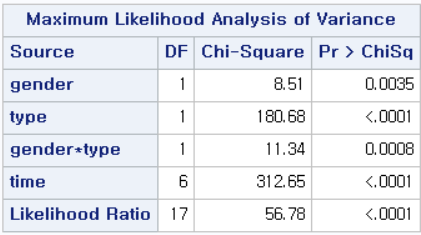
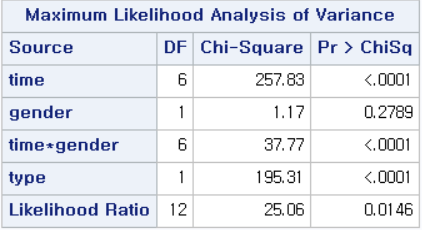
(XZ, YZ) 모형의 경우, =47.33-5.78=41.55 > =12.56

이므로, 유의수준 =0.05에서 (XZ, YZ) 모델이 더 적합하다고 말하기 힘들다

(XY, YZ) 모형의 경우, =12.74-5.78=6.96 < =12.56

이므로, 유의수준 =0.05에서 (XY, YZ) 모델이 더 적합하다고 할 수 있다.

따라서 (XY, YZ) 모델을 따르고, 여기서 더 축소된 모델들을 비교한다.



(XY, Z) 모형의 경우, =25.06-12.74=12.32 > =3.841

이므로, 유의수준 =0.05에서 (XY, Z) 모델이 더 적합하다고 말하기 힘들다

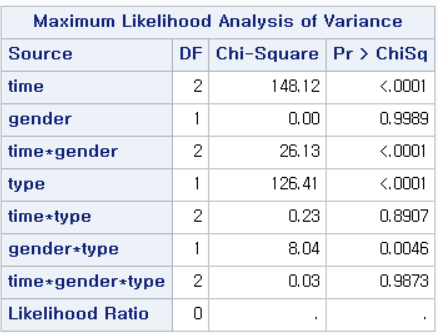
(YZ, X) 모형의 경우, =56.78-12.74=44.04 > =12.56

이므로, 유의수준 =0.05에서 (YZ, X) 모델이 더 적합하다고 말하기 힘들다

따라서 우리의 최종 모형은 (XY, YZ) 즉 (시간\*성별, 성별\*학교 계열) 이다.

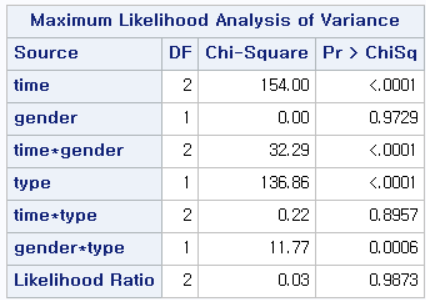
(2)

시간을 크게, 30분 미만 / 30분~2시간 / 2시간 이상 의 세 구간으로 나눈 뒤, (1)과 같은 과정을 반복하였다.

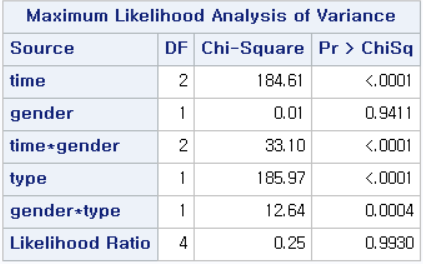
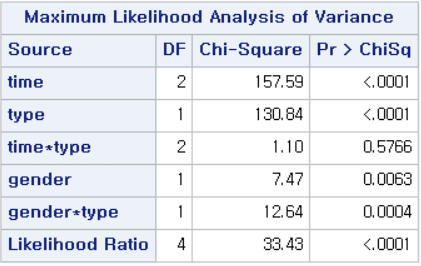
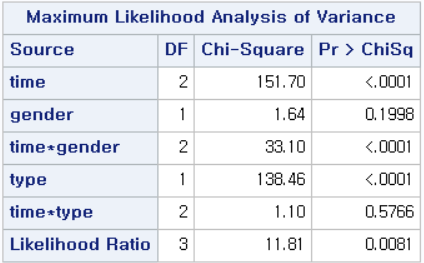


먼저 포화 모형을 고려한다.

3요인 교호작용에 대한 p-value는 0.9873으로 0.05보다 크기 때문에 유의수준 =0.05에서, 3요인 교호작용이 그리 유의하지 못함을 알 수 있다. 따라서 3요인 교호작용을 제외한 부분 독립 모형을 고려한다.



Likelihood Ratio의 카이스퀘어 통계량의 P-value가 0.9873으로, 유의수준 =0.05에서 모델이 옳게 적합 되었다고 이야기 할 수 있다. 이것이 최종 모형인 것인가를 알아보기 위해 더 축소된 모델들과 비교를 계속한다.



.

.= 우도비 검정 통계량.

=- : 두 모형의 비교를 위한 통계량

X=시간, Y=성별, Z=고등학교 종류 라고 할 때,

(XY, XZ) 모형의 경우, =11.81-0.03=11.78 > =3.841

이므로, 유의수준 =0.05에서 (XY, XZ) 모델이 더 적합하다고 말하기 힘들다

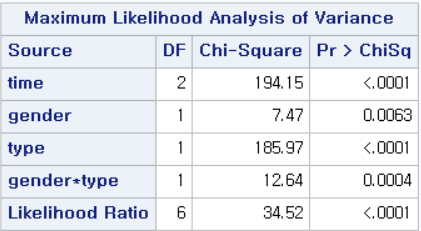
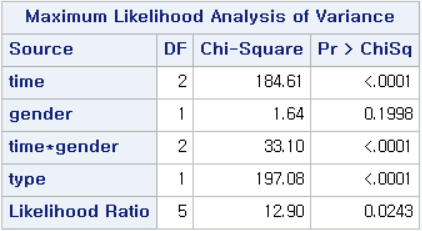
(XZ, YZ) 모형의 경우, =33.43-0.03=33.40 > =5.99

이므로, 유의수준 =0.05에서 (XZ, YZ) 모델이 더 적합하다고 말하기 힘들다

(XY, YZ) 모형의 경우, =0.25-0.03=0.22 < =5.99

이므로, 유의수준 =0.05에서 (XY, YZ) 모델이 더 적합하다고 할 수 있다.

따라서 (XY, YZ) 모델을 따르고, 여기서 더 축소된 모델들을 비교한다.



(XY, Z) 모형의 경우, =12.90-0.25=12.65 > =3.841

이므로, 유의수준 =0.05에서 (XY, Z) 모델이 더 적합하다고 말하기 힘들다

(YZ, X) 모형의 경우, =34.52-0.25=34.27 > =5.99

이므로, 유의수준 =0.05에서 (YZ, X) 모델이 더 적합하다고 말하기 힘들다

따라서 우리의 최종 모형은 (XY, YZ) 즉 (시간\*성별, 성별\*학교 계열) 이다.

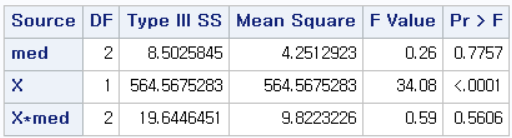
이는 (1)에서 가장 적절한 모형으로 선택한 것과 같다.

다만, (2)에서처럼, 범주를 3개로 줄인 것의 적합성에 대한 p-value가 훨씬 높아서, 두 번째 경우가 적합이 더 잘 된 것이라고 할 수 있다.

5.3

(1)

공분산 분석을 사용하기 위해서는 치료 방법과 실험 전 나병균의 수 간에 교호작용이 없어야 하므로 먼저 이를 검정한다.

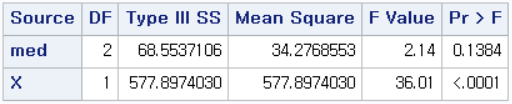


치료 방법과 나병균 수 간의 교호작용을 고려한 모델에서, 두 변수간 교호작용에 대한 F 통계량의 p-value가 0.5606으로 매우 높기 때문에, 유의수준 =0.05 하에서 두 변수간 교호작용이 유의하지 않다고 할 수 있다.

따라서 공분산 분석을 사용할 수 있으며, 실험 전 나병균 수를 제어하고 실시한 공분산 분석 결과는 다음과 같다.

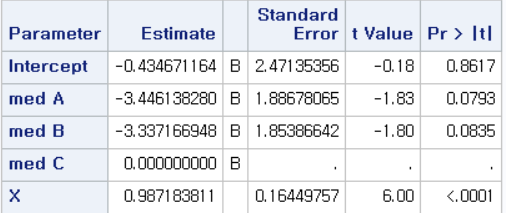
(i = 1,2,3 / j = 1, … , 10)

( 치료 방법 별 효과 / = 실험 전 나병균 수의 효과)



공분산 분석을 실시한 결과, 치료 방법에 대한 F-value가 2.14이며, p-value가 0.1384이다. 즉, 유의수준 =0.05 하에서 치료 방법간에 치료 후의 나병균 수에 미치는 영향의 차이가 유의하지 않다고 할 수 있다.

(2) 분석 결과를 토대로 보정된 평균은,



0.434671164 - 3.446138280 + 0.987183811\*10.73333 = 6.7149635

0.434671164 - 3.337166948 + 0.987183811\*10.73333 = 6.8239348

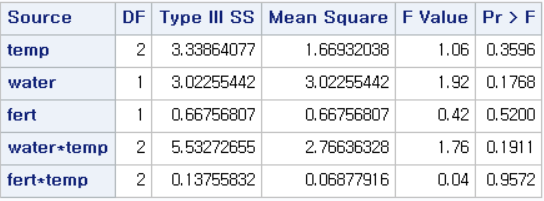
0.434671164 - 0 + 0.987183811\*10.73333 = 10.1611017

치료약 A가 약 6.7150, 치료약 B가 약 6.8240, 치료약 C가 약 10.1611 정도이다.

5.5

(1)

공분산 분석을 사용하기 위해서는 온도와 물의 양, 온도와 비료의 양 간에 각각 교호작용이 없어야 하므로 먼저 이를 검정한다.



온도와 물의 양, 온도와 비료의 양 간의 교호작용을 모두 고려한 모델에서,

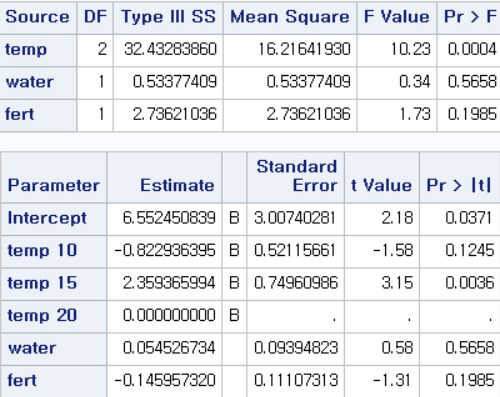
우선, 온도와 물의 양 간 교호작용에 대한 F 통계량의 p-value가 0.1911으로 높기 때문에, 유의수준 =0.05 하에서 두 변수간 교호작용이 유의하지 않다고 할 수 있다.

또한, 온도와 비료의 양 간 교호작용에 대한 F 통계량의 p-value가 0.9572로 높기 때문에, 유의수준 =0.05 하에서 두 변수간 교호작용이 유의하지 않다고 할 수 있다.

따라서 공분산 분석을 사용할 수 있으며, 물의 양과 비료의 양을 공변량으로 하는 공분산 분석은 다음과 같다.

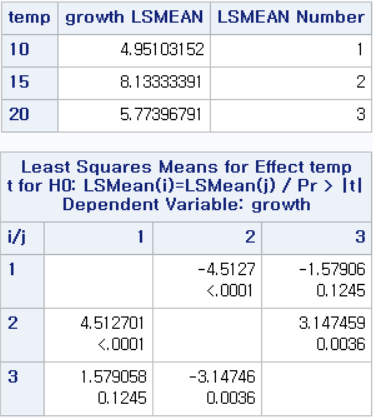
(i = 1,2,3 / j = 1, … , 12)

( 물 온도 별 효과 / = 물의 양의 효과 / = 비료량의 효과)



공분산 분석을 실시한 결과, 온도에 대한 F-value가 10.23이며, p-value가 0.0004이다. 즉, 유의수준 =0.05 하에서 온도가 박하의 성장속도에 미치는 영향의 차이가 유의 하다고 할 수 있다.

(2)



- 0.822936395 + 0.054526734\*24.9083333-0.145957320\*14.6388889 = 4.9510

+ 2.359365994 + 0.054526734\*24.9083333-0.145957320\*14.6388889 = 8.1333

+ 0 + 0.054526734\*24.9083333-0.145957320\*14.6388889 = 5.7740

이며,

,

,

,

를 검정한다. 검정에는 Bonferroni 방법을 사용하며 유의수준은 =0.05를 적용한다. 따라서 각각의 검정의 유의수준은 = (0.05/3)=0.16667이 된다.

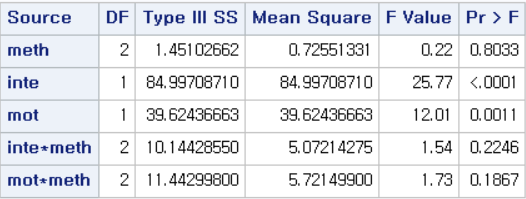
1. 의 경우, t통계량의 p-value가 0.0001, 3)의 경우 0.0036으로,

각각 유의수준 = (0.05/3)=0.016667에서 귀무가설을 기각하므로, 온도의 두 수준에 따라서 성장속도에는 유의한 차이가 있다고 이야기 할 수 있다.

2)의 경우 0.1245로 유의수준 = (0.05/3)=0.016667에서 귀무가설을 기각하지 못하므로, 온도의 두 수준에 따라서 성장 속도에 유의한 차이가 있다고 이야기 하기 어렵다

5.7

(1)



수업 방법과 지능, 수업 방법과 동기 간의 교호작용을 모두 고려한 모델에서,

우선, 수업 방법과 지능 간 교호작용에 대한 F 통계량의 p-value가 0.2246으로 높기 때문에, 유의수준 =0.05 하에서 두 변수간 교호작용이 유의하지 않다고 할 수 있다.

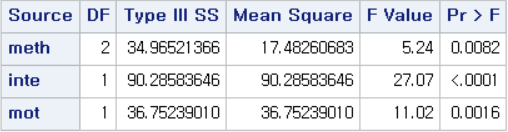
또한, 수업 방법과 동기 간 교호작용에 대한 F 통계량의 p-value가 0.1867로 높기 때문에, 유의수준 =0.05 하에서 두 변수간 교호작용이 유의하지 않다고 할 수 있다.

(2)

교호작용이 없으므로 공분산 분석으로 적합을 시킨다.

(i = 1,2,3 / j = 1, … , 20)

( 수업 방법 별 효과 / = 지능의 효과 / = 동기의 효과)



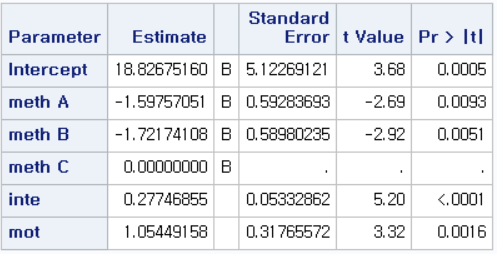
공분산 분석을 실시한 결과, 수업 방법에 대한 F-value가 5.24이며, p-value가 0.0082이다. 즉, 유의수준 =0.05 하에서 수업 방법이 성취에 미치는 영향의 차이가 유의 하다고 할 수 있다.

지능에 대한 F 통계량의 p-value는 0.0001 미만이다. 즉, 유의수준 =0.05 하에서 지능이 성취에 미치는 영향은 유의 하다고 할 수 있다.

동기에 대한 F 통계량의 p-value는 0.0016 이다. 즉, 유의수준 =0.05 하에서 동기가 성취에 미치는 영향은 유의 하다고 할 수 있다.

(3)

위를 통하여 공분산 분석을 통하여 수업 방법 별로 모형식을 쓰면 다음과 같다.



: 전통적 수업 방법

: 수준별 수업 방법

: 혼합형 수업 방법

/\*4-1\*/

**DATA** Guilt;

INPUT reside $ arrest $ imprison $ obs @@;

CARDS;

N Y Y 42

N Y N 109

N N Y 17

N N N 75

Y Y Y 33

Y Y N 175

Y N Y 53

Y N N 359

;

**PROC** **LOGISTIC** DATA= Guilt DESCENDING;

FREQ obs;

CLASS reside arrest imprison;

MODEL imprison = reside arrest/SCALE = NONE AGGREGATE;

**RUN**;

/\*4-5\*/

**DATA** psycho;

INPUT degree class incident @@;

CARDS;

4 1 1 4 1 9 4 1 4 4 1 3 4 0 2

4 1 0 4 0 1 4 1 3 4 1 3 4 1 7

4 0 1 4 0 2 3 1 5 3 0 6 3 1 3

3 0 1 3 1 8 3 1 2 3 0 5 3 1 5

3 1 9 3 0 3 3 1 3 3 1 1 2 0 0

2 1 4 2 0 3 2 0 9 2 1 6 2 0 4

2 0 3 1 1 8 1 1 2 1 1 7 1 0 5

1 0 4 1 0 4 1 1 8 1 0 8 1 0 9

;

**PROC** **LOGISTIC** DATA=psycho ORDER=DATA;

CLASS class;

MODEL degree = class | incident/SCALE=NONE AGGREGATE;

OUTPUT OUT=prob pred=p;

**RUN**;

**PROC** **LOGISTIC** DATA=psycho ORDER=DATA;

CLASS class;

MODEL degree = class incident/SCALE=NONE AGGREGATE;

OUTPUT OUT=prob pred=p;

**RUN**;

**PROC** **PRINT** data=prob(obs=**10**);

**RUN**;

/\*4-7\*/

**DATA** safety;

INPUT seatbelt $ bounce $ death $ obs @@;

CARDS;

Y Y Y 659 Y Y N 270 Y N Y 532 Y N N 347

N Y Y 432 N Y N 532 N N Y 269 N N N 552

;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt | bounce | death;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt|bounce seatbelt|death bounce|death;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt|bounce seatbelt|death;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt|bounce bounce|death;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt|death bounce|death;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN death seatbelt|bounce;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt|death bounce;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN bounce|death seatbelt;

**RUN**;

**PROC** **CATMOD** DATA = safety;

WEIGHT obs;

MODEL seatbelt\*bounce\*death = \_response\_/ NORESPONSE NOITER;

LOGLIN seatbelt bounce death;

**RUN**;

**PROC** **LOGISTIC** DATA=safety;

FREQ obs;

CLASS seatbelt bounce death;

MODEL death = seatbelt | bounce/SCALE=NONE AGGREGATE;

**RUN**;

**PROC** **LOGISTIC** DATA=safety DESCENDING;

FREQ obs;

CLASS seatbelt bounce death;

MODEL death = seatbelt bounce/SCALE =NONE AGGREGATE;

**RUN**;

/\*4-9\*/

**DATA** reading;

INPUT time $ gender $ type $ obs @@;

CARDS;

<0m m arts 34 <0m m work 64 <0m w arts 49 <0m w work 135

<30m m arts 29 <30m m work 61 <30m w arts 31 <30m w work 118

<60m m arts 40 <60m m work 81 <60m w arts 37 <60m w work 142

<120m m arts 37 <120m m work 65 <120m w arts 32 <120m w work 64

<240m m arts 24 <240m m work 40 <240m w arts 11 <240m w work 37

<360m m arts 6 <360m m work 15 <360m w arts 0 <360m w work 3

>360m m arts 3 >360m m work 7 >360m w arts 4 >360m w work 2

;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time | gender | type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time|gender time|type gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time|gender gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time|gender time|type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time|type gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time|gender type;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN gender|type time;

**RUN**;

**PROC** **CATMOD** DATA = reading;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/NOITER PRED=FREQ;

LOGLIN time gender|type;

**RUN**;

/\*4-9-2\*/

**DATA** reading2;

INPUT time $ gender $ type $ obs @@;

CARDS;

<0.5h m arts 63 <0.5h m work 125 <0.5h w arts 80 <0.5h w work 253

<2h m arts 77 <2h m work 146 <2h w arts 69 <2h w work 206

>2h m arts 33 >2h m work 62 >2h w arts 15 >2h w work 42

;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time | gender | type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time|gender time|type gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time|gender gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time|gender time|type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time|type gender|type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time|gender type;

**RUN**;

**PROC** **CATMOD** DATA = reading2;

WEIGHT obs;

MODEL time\*gender\*type = \_response\_/ NORESPONSE NOITER;

LOGLIN time gender|type;

**RUN**;

/\*DATA EXPORT FOR R\*/

**proc** **export** data=Guilt

outfile='C:\Users\pkmon\Desktop\Guilt.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=psycho

outfile='C:\Users\pkmon\Desktop\psycho.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=safety

outfile='C:\Users\pkmon\Desktop\safety.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=reading

outfile='C:\Users\pkmon\Desktop\reading.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=reading2

outfile='C:\Users\pkmon\Desktop\reading2.csv'

dbms=csv

replace;

**run**;

/\*5-3\*/

**DATA** na;

INPUT med $ X Y @@;

CARDS;

A 11 6 A 8 0 A 5 2 A 14 8 A 19 11

A 6 4 A 10 13 A 6 1 A 11 8 A 3 0

B 6 0 B 6 2 B 7 3 B 8 1 B 18 18

B 8 4 B 19 14 B 8 9 B 5 1 B 15 9

C 16 13 C 13 10 C 11 18 C 9 5 C 21 23

C 16 12 C 12 5 C 12 16 C 7 1 C 12 20

;

**PROC** **GLM** DATA=na;

CLASS med;

MODEL Y = med X X\*med /SOLUTION;

LSMEANS med/TDIFF;

**RUN**;

**PROC** **GLM** DATA=na;

CLASS med;

MODEL Y = med X/SOLUTION;

LSMEANS med/TDIFF;

**RUN**;

**PROC** **MEANS** DATA=na;

VAR X;

**RUN**;

/\*5-5\*/

**DATA** baka;

INPUT temp water fert growth @@;

CARDS;

10 23.5 15 5 10 24 15.5 4.5 10 23 14 5 10 24.5 13.5 4.5 10 22.5 13.5 5.5 10 24.5 13.5 3

10 25.5 13 5 10 25 14 5 10 23 14.5 5.5 10 23 14 4.5 10 22.5 14 6.5 10 23 15 5.5

15 28.5 16 7 15 26 15.5 9 15 28.5 17 8.5 15 26.5 16 8.5 15 27 18.5 6 15 28 14.5 8

15 26.5 17.5 7 15 29 13.5 7 15 27 16.5 7 15 27 16.5 11 15 27 17.5 9 15 27 18.5 8

20 25.6 12 6.6 20 23.1 11.6 9.2 20 20.8 18.2 5.6 20 19.2 11.9 4.2 20 21.3 20.2 4.6 20 26.6 13.1 4.5

20 25.5 16.2 4.5 20 18.9 11.9 6.1 20 29.3 10.5 6.3 20 30.2 11.3 8.1 20 18.9 10.9 4.9 20 25.3 12.2 6.2

;

**PROC** **GLM** DATA = baka;

CLASS temp;

MODEL growth = temp water fert water\*temp fert\*temp;

**RUN**;

**PROC** **GLM** DATA = baka;

CLASS temp;

MODEL growth = temp water fert/SOLUTION;

LSMEANS temp/TDIFF;

**RUN**;

**PROC** **MEANS** DATA=baka;

VAR water fert;

**RUN**;

/\*5-7\*/

**DATA** study;

INPUT meth $ inte mot ach @@;

CARDS;

A 99 12 58 A 95 11 55 A 99 13 59 A 99 13 55 A 102 12 60 A 101 12 57

A 102 12 57 A 103 13 61 A 102 13 58 A 103 14 59 A 102 14 62 A 104 13 59

A 106 15 60 A 106 13 62 A 107 13 60 A 108 12 62 A 109 13 64 A 111 13 63

A 110 13 59 A 113 13 63

B 95 12 58 B 97 12 55 B 100 13 60 B 100 13 55 B 102 13 60 B 105 12 58

B 102 13 57 B 102 14 62 B 105 12 58 B 105 13 61 B 106 14 63 B 109 13 59

B 107 14 61 B 108 14 64 B 111 14 61 B 108 13 63 B 111 13 65 B 113 14 64

B 114 13 62 B 114 14 60

C 94 12 59 C 95 12 56 C 99 13 62 C 99 11 57 C 99 14 61 C 102 12 60 C 99 13 58

C 101 13 63 C 104 13 60 C 101 14 63 C 104 15 65 C 107 13 61 C 107 13 62

C 106 14 64 C 109 14 61 C 107 14 65 C 109 14 63 C 111 14 66 C 111 14 62

C 112 13 64

;

**PROC** **GLM** DATA=study;

CLASS meth;

MODEL ach = meth inte mot meth\*inte meth\*mot;

MODEL ach = meth inte mot/SOLUTION;

**RUN**;

**PROC** **GLM** DATA=study;

CLASS meth;

MODEL ach = meth inte mot/SOLUTION;

LSMEANS meth/TDIFF;

**RUN**;

**proc** **export** data=na

outfile='C:\Users\pkmon\Desktop\생과데 과제3\na.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=baka

outfile='C:\Users\pkmon\Desktop\생과데 과제3\baka.csv'

dbms=csv

replace;

**run**;

**proc** **export** data=study

outfile='C:\Users\pkmon\Desktop\생과데 과제3\study.csv'

dbms=csv

replace;

**run**;